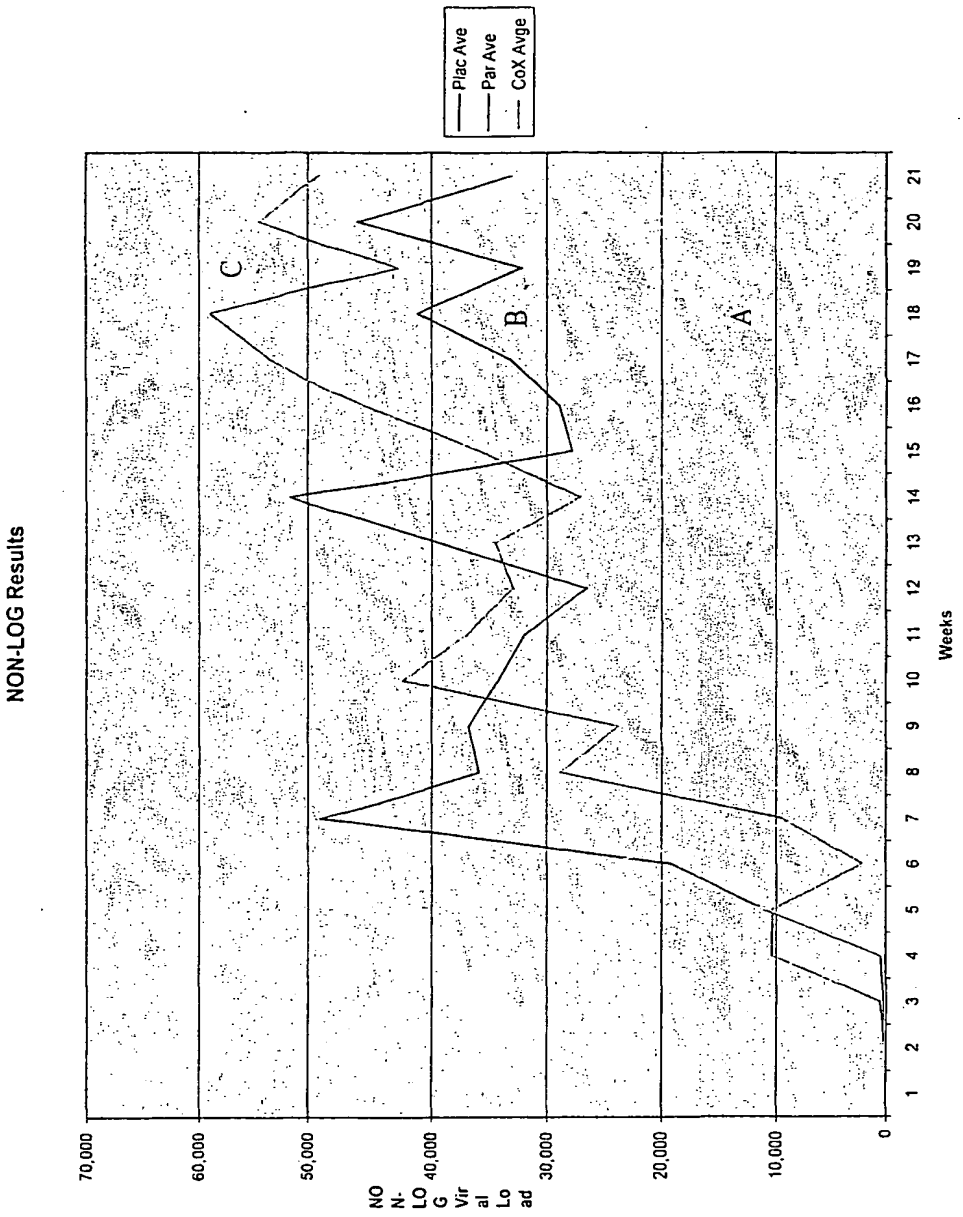

Figure 1

BEST AVAILABLE COPY

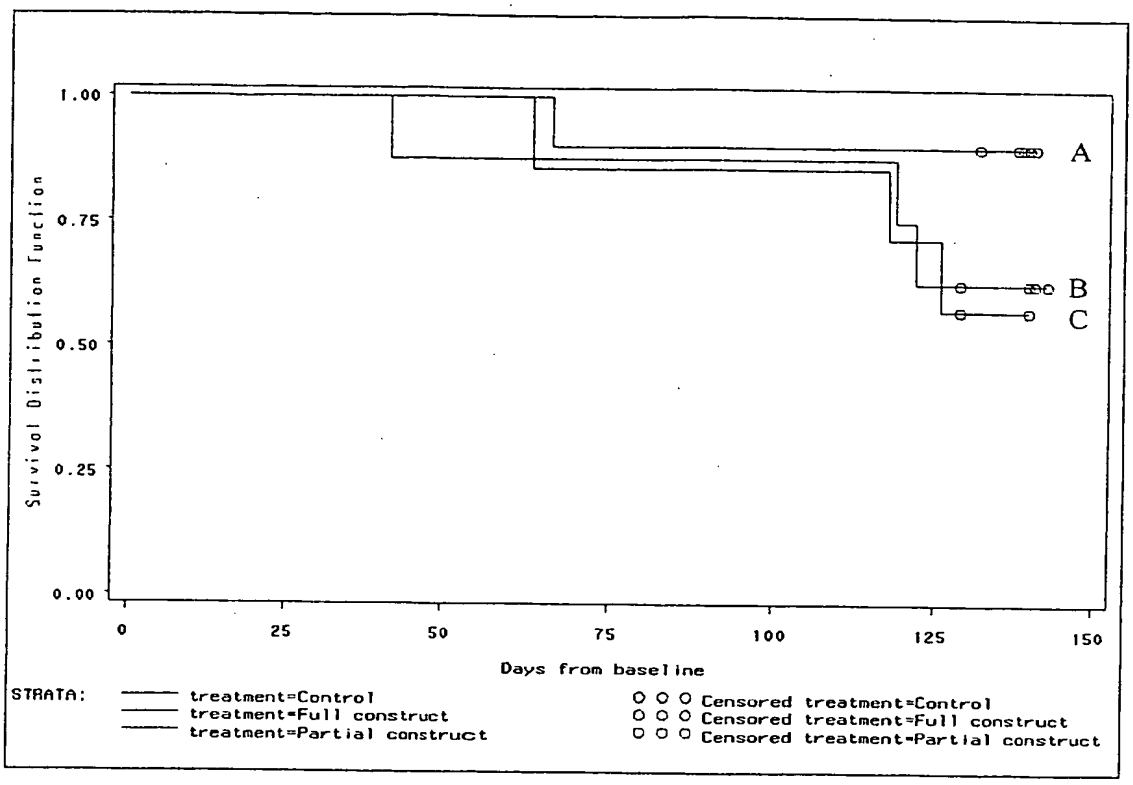
1.00
0.75
0.50
0.25
0.00
Survival Distribution Function
0
25
50
75
100
125
150
Days from baseline
A
B
C
STRATA:
treatment=Control
treatment=Full construct
treatment=Partial construct
Censored treatment=Control
Censored treatment=Full construct
Censored treatment=Partial construct

FIGURE 2

DNA sequence of the insertion site of VIR201 containing HIV gag/pol, human interferon and reporter cassette (Ecogpt & beta-galactosidase) inserts

```
   1 AGACAGTTATCCCAATACGGTATACAAGGAGACAATTTATCAATTTTTGTAGATTCTTCC
     TCTGTCAATAGGGTTATGCCATATGTTCCTCTGTTAAATAGTTAAAAACATCTAAGAAGG
Fowlpox virus 5' flanking region of insertion site -->
  61 AATGAAGTTGCTATAAACAGGCACTCTATTATAGGAGCTAGACAGTTGAATCCTATATGC
     TTACTTCAACGATATTTGTCCGTGAGATAATATCCTCGATCTGTCAACTTAGGATATACG
 121 GTAGTATCTTTTTATCCCTTTGATCCAGAACATAAAGTTTTTTTCGTTATATATGTTGGT
     CATCATAGAAAAATAGGGAAACTAGGTCTTGTATTTCAAAAAAAGCAATATATACAACCA
 181 AGATATAAAGATAAGTATTGTGGAATTTCCTACGTAGCTGATAGAGAAGATATGTACAAA
     TCTATATTTCTATTCATAACACCTTAAAGGATGCATCGACTATCTCTTCTATACATGTTT
 241 GTTATCAACAGGATATACCCGTACGTTAGTTGTTTTTACCTCGTATCAGATGGTATAATA
     CAATAGTTGTCCTATATGGGCATGCAATCAACAAAAATGGAGCATAGTCTACCATATTAT
 301 AATTTTCATACTACTCCCGTAGCTAATCACACTAGAAATATTAAACCCCTTCCAGTTAAT
     TTAAAAGTATGATGAGGGCATCGATTAGTGTGATCTTTATAATTTGGGGAAGGTCAATTA
 361 TATTGTAATACTTTATGTGAAATAGTATATGATTTTGAATATTTAAAGTTTGAACAAGGT
     ATAACATTATGAAATACACTTTATCATATACTAAACTTATAAATTTCAAACTTGTTCCA
 421 GTTATGTCTATTCCGGTGTTCATGCCTTTTGTACCAAAACAGTTTGTATCTATTATCAAT
     CAATACAGATAAGGCCACAAGTACGGAAAACATGGTTTTGTCAAACATAGATAATAGTTA
```

Figure 3 continued

```
481  TTACCAGATGATATTCTCATAACATGTACAGCGTCCAGTAACATAGAATACATAACACAT
     AATGGTCTACTATAAGAGTATTGTACATGTCGCAGGTCATTGTATCTTATGTATTGTGTA

541  ATAGATAATAAAAAGCTAAAAAGAATACTTATAATAAAAGATAAATTTCTAAAGGGT
     TATCTATTATTTTTCGATTTTTCTTATGAATATTATTTATTTCTATTTAAGATTTCCCA

601  ACTATCATGCAAGGTACTTTTAAAAAGTAAATATCATAAGACACAAGAAGTATACATAT
     TGATAGTACGTTCCATGAAAATTTTTTCATTTATAGTATTCTGTGTTCTTCATATGTATA

661  ACTATAACGTATTCTTCTTTTGATTGCCCTAAACTAGAAGATACTAAGTCATCGCTGCCA
     TGATATTGCATAAGAAGAAAACTAACGGGATTTGATCTTCTATGATTCAGTAGCGACGGT

721  AGTACGTGCAATAAAGCCATATTAGATGGGCGTAGATATGTTACAAAAACTTTTAATGAT
     TCATGCACGTTATTTCGGTATAATCTACCCGCATCTATACAATGTTTTTGAAAATTACTA

781  ACAATATAAATGGAAATAGCTAGAGAAACGCTAATAACGATAGGCCTTACTATATTAGTA
     TGTTATATTTACCTTTATCGATCTCTTTGCTATTATTGCTATCCGGAATGATATAATCAT

841  GTGTTATTGATAATAACTGGATTCTCGCTAGTGCTAAGATTAATACCGGGTGTTTATAGT
     CACAATAACTATTATTGACCTAAGAGCGATCACGATTCTAATTATGGCCCACAAATATCA

901  TCAGTATCGAGGTCATCATTTACAGCAGGAAGAATACTTCGTTTTATGGAAATATTTTCT
     AGTCATAGCTCCAGTAGTAAATGTCGTCCTTCTTATGAAGCAAAATACCTTTATAAAAGA

961  ACTATTATGTTTATTCCTGGAATAATTATATTGTACGCTGCTTATATAAGAAAAATTAAA
     TGATAATACAAATAAGGACCTTATTAATATAACATGCGACGAATATATTCTTTTTAATTT

1021 ATGAAAAATAATTAGAATCTGAAAATGTCTTCTGGAAGCATCCATGTTATTACAGGCCCT
     TACTTTTTATTAATCTTAGACTTTTACAGAAGACCTTCGTAGGTACAATAATGTCCGGGA
```

Figure 3 continued

```
                                                    >M S S G S I H V I T G P
                                                    Fowlpox virus tk protein coding sequence→

1081 ATGTTTTCCGGTAAAACATCGGAGCTAGTAAGAAGAATAAAAAGATTTATGCTATCTAAC
       TACAAAAGGCCATTTTGTAGCCTCGATCATTCTTCTTATTTTTCTAAATACGATAGATTG
         >M F S G K T S E L V R R I K R F M L S N

1141 TTTAAATGTATTATTATTAAACATTGTGGAGATAAATAGATATAATGAGGATGATATAAAC
       AAATTTACATAATAATAATTTGTAACACCTCTATTATCTATATTACTCCTACTATATTTG
         >F K C I I I K H C G D N R Y N E D D I N

1201 AAAGTATATACTCATGATCTATTGTTTATGGAGGCTACGGCATCTTCTAATCTATCTGTA
       TTTCATATATGAGTACTAGATAACAAATACCTCCGATGCCGTAGAAGATTAGATAGACAT
         >K V Y T H D L L F M E A T A S S N L S V

1261 TTAGTACCTACGCTATTAAATGATGGAGTTCAGGTAATAGGTATAGACGAGGCTCAATTC
       AATCATGGATGCGATAATTTACTACCTCAAGTCCATTATCCATATCTGCTCCGAGTTAAG
         >L V P T L L N D G V Q V I G I D E A Q F

1321 TTTCTAGACATAGTAGAATTTAGTGAATCCATGGCTAATTTAGGTAAAACAGTTATTGTG
       AAAGATCTGTATCATCTTAAATCACTTAGGTACCGATTAAATCCATTTTGTCAATAACAC
         >F L D I V E F S E S M A N L G K T V I V

1381 GCCGCGCTTAACGGTGATTTTAAACGCGAATTATTCGGTAACGTATATAAGTTATTATCA
       CGGCGCGAATTGCCACTAAAATTTGCGCTTAATAAGCCATTGCATATATTCAATAATAGT
         >A A L N G D F K R E L F G N V Y K L L S

1441 TTAGCTGAAACAGTGTCCAGTTTGACAGCTATTTGCGTGAAATGCTATTGCGACGCTTCG
       AATCGACTTTGTCACAGGTCAAACTGTCGATAAACGCACTTTACGATAACGCTGCGAAGC
```

Figure 3 continued

```
      > L  A  E  T  V  S  S  S  L  T  A  I  C  V  K  C  Y  C  D  A  S

1501  TTTTCTAAACGAGTTACAGAAAATAAAGAAGTAATGGATATAGGTGGTAAAGATAAATAC
        AAAAGATTTGCTCAATGTCTTTTATTTCTTCATTACCTATATCCACCATTTCTATTTATG
      > F  S  K  R  V  T  E  N  K  E  V  M  D  I  G  G  K  D  K  Y

1561  ATAGCCGTGTGTAGGAAATGTTTTTTAGTAATTAAgggagatctccccatggcccaaa
        TATCGGCACACATCCTTTACAAAAAATCATTAATTccccctctagaggggtaccgggttt
      > I  A  V  C  R  K  C  F  F  S  N  •

1621  gcggggtttgaacagggtttcgctcaggtttgcctgtgtcatggatgcagcctccagaat
        cgccccaaacttgtcccaaagcgagtccaaacggacacagtacctacgtcggaggtctta

1681  acttactggaaactattgtaacccgcctgaagttaaaaagaacaacgcccggcagtgcca
        tgaatgacctttgataacattgggcggacttcaatttttcttgttgcgggccgtcacggt

1741  ggcgttgaaaagaTTAGCGACCGGAGATTGGCGGGACGAATACGACGCCCATATCCCACG
        ccgcaacttttctAATCGCTGGCCTCTAACCGCCCTGCTTATGCTGCGGGTATAGGGTGC
            < •  R  G  S  I  P  P  V  F  V  V  G  M  D  W  P
                    End of Ecogpt protein coding sequence

1801  GCTGTTCAATCCAGGTATCTTGCGGGATATCAACACATAGTCATCAACCAGCGGACGAC
        CGACAAGTTAGGTCCATAGAACGCCCTATAGTTGTTGTATCAGTAGTTGGTCGCCTGCTG
      < Q  E  I  W  T  D  Q  P  I  D  V  V  Y  D  D  V  L  P  R  G

1861  CAGCCGGTTTTGCGAAGATGGTGACAAAGTGCGCTTTTGGATACATTTCACGAATCGCAA
        GTCGGCCAAAACGCTTCTACCACTGTTTCACGCGAAAACCTATGTAAAGTGCTTAGCGTT
      < A  P  K  A  F  I  T  V  F  H  A  K  P  Y  M  E  R  I  A  V
```

Figure 3 continued

```
1921 CCGCAGTACCACCGGTATCCACCAGGTCATCAATAACGATGAAGCCTTCGCCATCGCCTT
       GGCGTCATGGTGGCCATAGGTGGTCCAGTAGTTATTGCTACTTCGGAAGCGGTAGCGGAA
         <  A  T  G  G  T  D  V  L  D  D  D  I  V  I  F  G  E  G  D  G  E

1981 CTGCGCGTTTCAGCACTTTAAGCTCGCGCTGGTTGTCGTGATCGTAGCTGGAAATACAAA
       GACGCGCAAAGTCGTGAAATTCGAGCGCGACCAACAGCACTAGCATCGACCTTTATGTTT
         <  A  R  K  L  V  K  L  E  R  Q  N  D  H  D  Y  S  S  I  C  V

2041 CGGTATCGACATGACGAATACCCAGTTCACGCGCCAGTAACGCACCCGGTACCAGACCGC
       GCCATAGCTGTACTGCTTATGGGTCAAGTGCGCGGTCATTGCGTGGGCCATGGTCTGGCG
         <  T  D  V  H  R  I  G  L  E  R  A  L  L  A  G  P  V  L  G  G

2101 CACGGCTTACGGCAATAATGCCTTTCCATTGTTCAGAAGGCATCAGTCGGCTTGCGAGTT
       GTGCCGAATGCCGTTATTACGGAAAGGTAACAAGTCTTCCGTAGTCAGCCGAACGCTCAA
         <  R  S  V  A  I  I  G  K  W  Q  E  S  P  M  L  R  S  A  L  K

2161 TACGTGCATGGATCTGCAACATGTCCCAGGTGACGATGTATTTTCGCTCATgtgaagtg
       ATGCACGTACCTAGACGTTGTACAGGGTCCACTGCTACATAAAAAGCGAGTAcacttcac
         <  R  A  H  I  Q  L  M  D  W  T  V  I  Y  K  E  S  M
                    ↖Start of Ecogpt protein coding sequence

2221 tcccagcctgtttatctacggcttaaaaagtgttcgagggaaaataggttgcgcgagat
       agggtcggacaaatagatgccgaatttttcacaagctcccctttatccaacgcgctcta

2281 tatagagatccgtcactgttctttatgatctacttcctaCCGTGCAATAAATTAGAATA
       atatctctaggcagtgacaagaaatactagatgaaggaatGGCACGTTATTTAATCTTAT

2341 TATTTTCTACTTTTACGAGAAATTAATTATTGTATTTATTATGGGTGAAAAACTTA
       ATAAAAGATGAAAATGCTCTTTAATTAATAACATAAATAAATACCCACTTTTTGAAT
```

Figure 3 continued

Figure 3 continued

```
2821 GGTGTGGGCCATAATTCAATTCGCGCGTCCCGCAGCGCAGACCGTTTTCGCTCGGGAAGA
     CCACACCCGGTATTAAGTTAAGCGCGCAGGGCGTCGCGTCTGGCAAAAGCGAGCCCTTCT
        <  H  P  G  Y  N  L  E  R  T  G  C  R  L  G  N  E  S  P  F  V

2881 CGTACGGGGTATACATGTCTGACAATGGCAGATCCCAGCGGTCAAAACAGGCGGCAGTAA
     GCATGCCCCATATGTACAGACTGTTACCGTCTAGGGTCGCCAGTTTTGTCCGCCGTCATT
        <  Y  P  T  Y  M  D  S  L  P  L  D  W  R  D  F  C  A  A  T  L

2941 GGCGGTCGGGATAGTTTCTTGCGGCCCTAATCCGAGCCAGTTTACCCGCTCTGCTACCT
     CCGCCAGCCCTATCAAAGAACGCCGGGATTAGGCTCGGTCAAATGGGCGAGACGATGGA
        <  R  D  P  Y  N  E  Q  P  G  G  L  G  L  W  N  V  R  E  A  V  Q

3001 GCGCCAGCTGGCAGTTCAGGCCAATCCGCGCCGGATGCGGTGTATCGCTCGCCACTTCAA
     CGCGGTCGACCGTCAAGTCCGGTTAGGCGCGGCCTACGCCACATAGCGAGCGGTGAAGTT
        <  A  L  Q  C  N  L  G  I  R  A  P  H  P  T  D  S  A  V  E  V

3061 CATCAACGGTAATCGCCATTTGACCACTACCATCAATCCGGTAGGTTTCCGGCTGATAA
     GTAGTTGCCATTAGCGGTAAACTGGTGATGGTAGTTAGGCCATCCAAAAGGCCGACTATT
        <  D  V  T  I  A  M  Q  G  S  G  D  I  R  Y  T  K  R  S  I  F

3121 ATAAGGTTTTCCCCTGATGCTGCCACGCGTGAGCGGTCGTAATCAGCACCGCATCAGCAA
     TATTCCAAAAGGGGACTACGACGGTGCGCACTCGCCAGCATTAGTCGTGGCGTAGTCGTT
        <  L  T  K  G  Q  H  Q  W  A  H  A  T  T  I  L  V  A  D  A  L

3181 GTGTATCTGCCGTGCACTGCAACAACGCTGCTTCGGCCTGGTAATGGCCCGCCGCCTTCC
     CACATAGACGGCACGTGACGTTGTTGCGACGAAGCCGGACCATTACCGGGCGGCGGAAGG
        <  T  D  A  T  C  Q  L  L  A  A  E  A  Q  Y  H  G  A  A  K  W

3241 AGCGTTCGACCCAGGCGTTAGGGTCAATGCGGGTCGCTTCACTTACGCCAATGTCGTTAT
```

Figure 3 continued

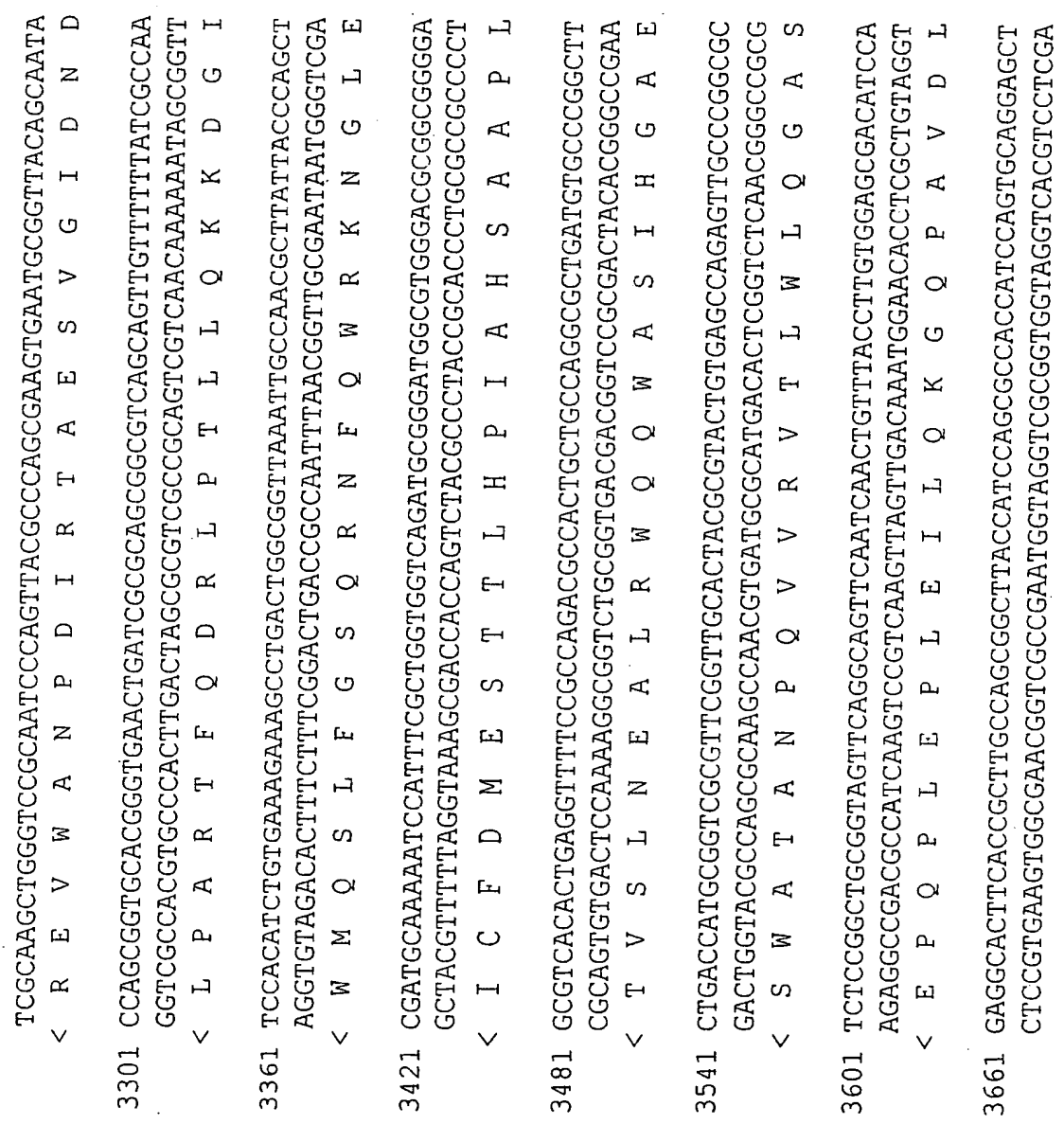

Figure 3 continued

```
                                                                  < P  V  E  G  G  S  A  L  P  K  K  G  D  D  L  A  A  V  M  W  W  H  H  L  L  E
                                                                  
3721 CGTTATCGCTATGACGGAACAGGTATTCGCTGGTCACTTCGATGGTTTGCCCGGATAAAC
       GCAATAGCGATACTGCCTTGTCCATAAGCGACCAGTGAAGCTACCAAACGGGCCTATTTG
          < N  D  S  H  R  F  L  L  Y  E  S  T  V  E  I  T  Q  G  S  L  R
                                                                  
3781 GGAACTGGAAAAACTGCTGGTGTTTTGCTTCCGTCAGCGCTGGATGCGGCGTGCGGT
       CCTTGACCTTTTTGACGACCACAAAACGAAGGCAGTCGCGACCTACGCCGCACGCCA
          < F  Q  F  F  Q  Q  Q  H  K  K  A  E  E  T  L  A  A  P  H  P  T  T  R  R  D
                                                                  
3841 CGGCAAAGACCAGACCGTTCATACAGAACTGGCGATCGTTCGGCGTATCGCCAAAATCAC
       GCCGTTTCTGGTCTGGCAAGTATGTCTTGACCGCTAGCAAGCCGCATAGCGGTTTTAGTG
          < A  F  V  L  G  N  M  C  F  Q  R  D  N  P  T  D  G  F  D  D  G
                                                                  
3901 CGCCGTAAGCCGACCACGGGTTGCCGTTTCATCATATTTAATCAGCGACTGATCCACCC
       GCGGCATTCGGCTGGTGCCCAACGGCAAAAGTAGTATAAATTAGTCGCTGACTAGGTGGG
          < G  Y  A  S  W  P  N  G  N  E  D  Y  K  K  I  L  L  S  Q  D  V  W
                                                                  
3961 AGTCCCAGACGAAGCCGCCCTGTAAACGGGGATACTGACGAAACGCCTGCCAGTATTTAG
       TCAGGGTCTGCTTCGGCGGGACATTTGCCCCTATGACTGCTTTGCGGACGGTCATAAATC
          < D  W  V  F  G  G  Q  L  R  P  Y  Q  R  F  A  Q  W  Y  K  A
                                                                  
4021 CGAAACCGCCAAGACTGTTACCCATCGCGTGGGCGTATTCGCAAAGGATCAGCGGGCGCG
       GCTTTGGCGGTTCTGACAATGGGTAGCGCACCCGCATAAGCGTTTCCTAGTCGCCCGCGC
          < F  G  G  L  S  N  G  M  A  H  A  Y  E  C  L  I  L  P  R  T
                                                                  
4081 TCTCTCCAGGTAGCGAAAGCCATTTTTTGATGGACCATTTCGGCACAGCCGGGAAGGGCT
       AGAGAGGTCCATCGCTTTCGGTAAAAAACTACCTGGTAAAGCCGTGTCGGCCCTTCCCGA
          < E  G  P  L  S  L  W  K  K  I  S  W  K  P  V  A  P  F  P  Q
```

Figure 3 continued

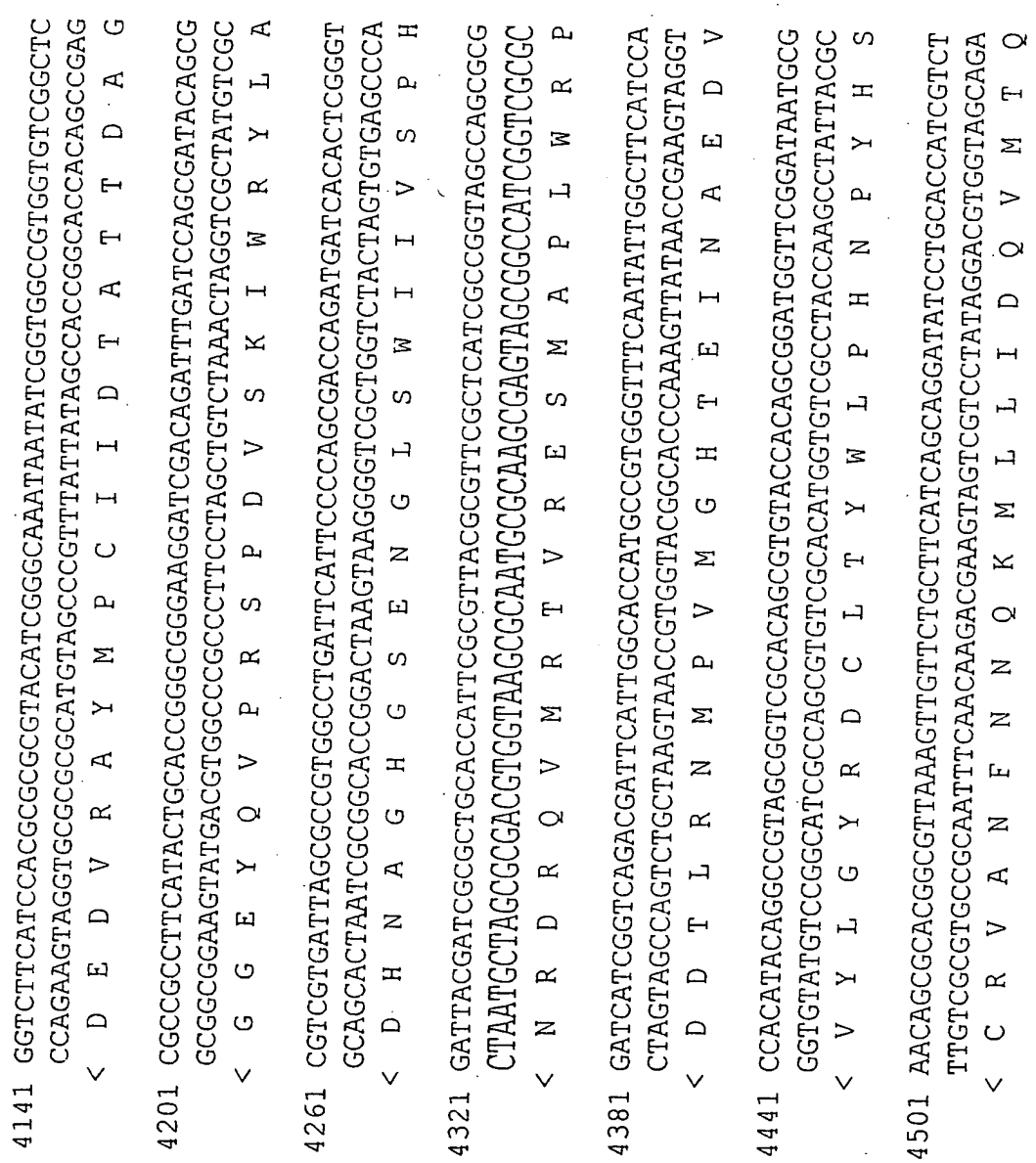

Figure 3 continued

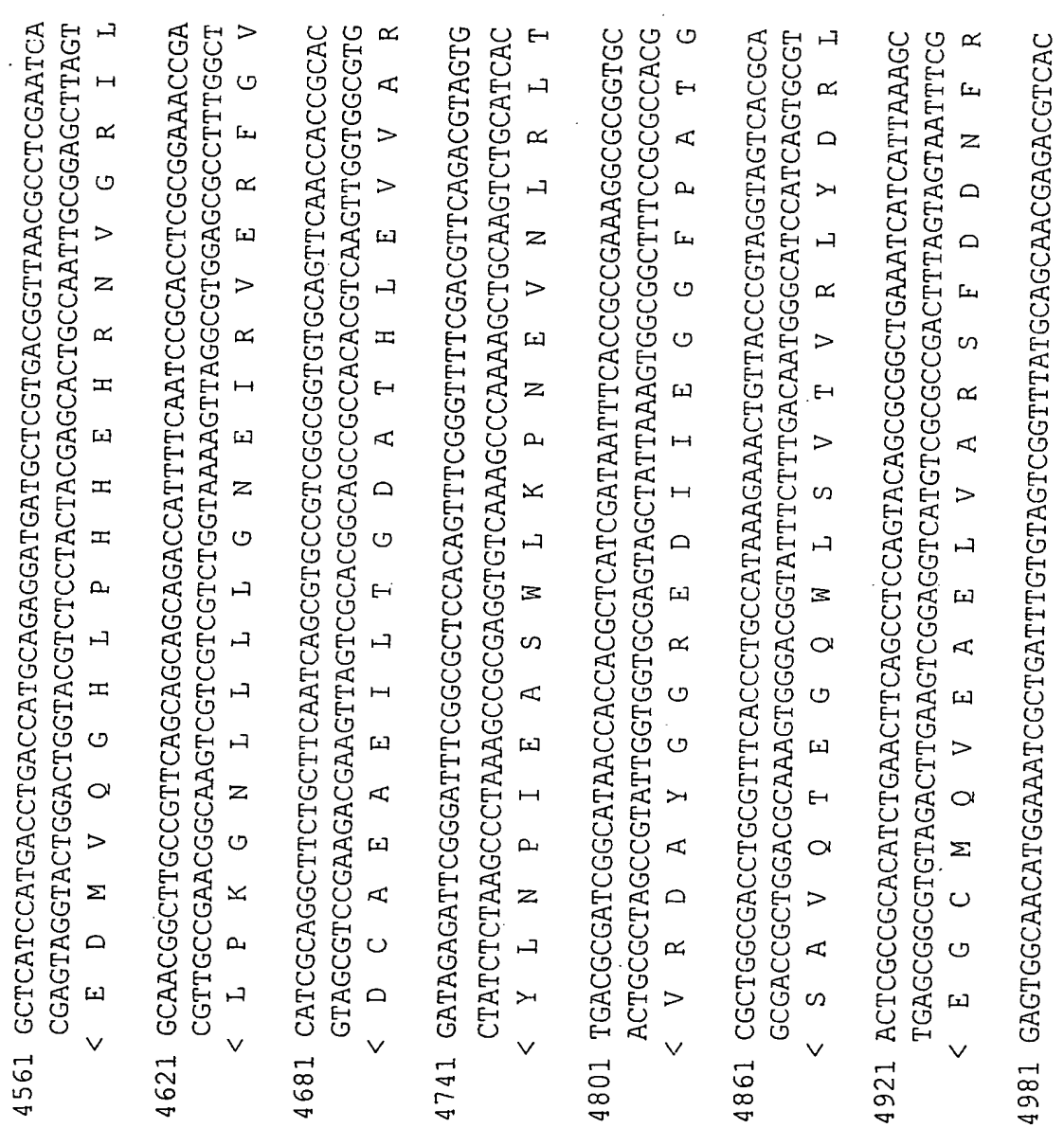

Figure 3 continued

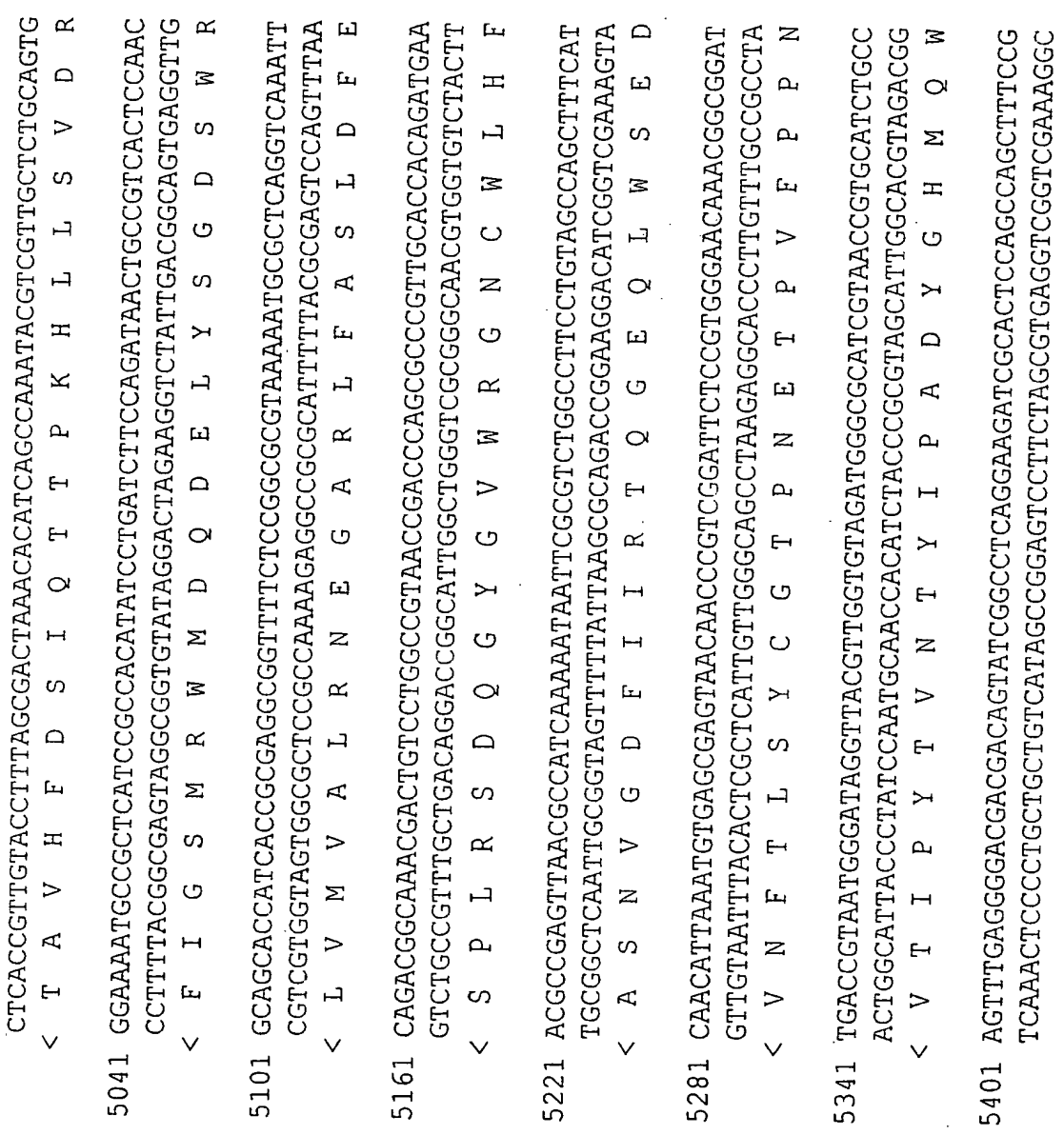

Figure 3 continued

```
      < N  S  P  V  V  V  T  D  A  E  P  L  D  D  C  E  L  W  S  E  P

5461  GCACCGCTTCTGGTGCCGGAAACCAGGCAAAGCGCCATTCGCCATTCAGGCTGCGCAACT
      CGTGGCGAAGACCACCGGCCTTTGGTCCGTTTCGCGGTAAGCGGTAAGTCCGACGCGTTGA
         < V  A  E  P  A  P  F  W  A  F  R  W  E  G  N  L  S  R  L  Q

5521  GTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGAT
      CAACCCTTCCCGCTAGCCACGCCCGGAGAAGCGATAATGCGGTCGACCGCTTTCCCCCTA
         < Q  S  P  R  D  T  R  A  E  E  S  N  R  W  S  A  F  P  P  H

5581  GTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAA
      CACGACGTTCCGCTAATTCAACCCATTGCGGTCCCAAAAGGGTCAGTGCTGCAACATTTT
         < A  A  L  R  N  L  Q  T  V  G  P  N  E  W  D  R  R  Q  L  V

5641  CGACGGGATCTAGCATggatctagccATTTAGTATCCTAAAATTGAATTGTAATTATCGA
      GCTGCCCTAGATCGTAcctagatcggTAAATCATAGGATTTTAACTTAACATTAATAGCT
         < V  P  D  L  M
                  ←Start of beta-Galactosidase protein coding sequence
                     ← Fowlpox virus bidirectional promoter (in bold) →

5701  TAATAAATGgacggatcgATGAAATATACAAGTTATATCTTGGCTTTTCAGCTCTGCATC
      ATTATTTACctgcctagcTACTTTATATGTTCAATATAGAACCGAAAAGTCGAGACGTAG
                  > M  K  Y  T  S  Y  I  L  A  F  Q  L  C  I
                              Human interferon gamma protein coding sequence→

5761  GTTTTGGGTTCTCTTGGCTGTTACTGCCAGGACCCATATGTAAAAGAAGCAGAAAACCTT
      CAAAACCCAAGAGAACCGACAATGACGGTCCTGGGTATACATTTTCTTCGTCTTTTGGAA
         > V  L  G  S  L  G  C  Y  C  Q  D  P  Y  V  K  E  A  E  N  L
```

Figure 3 continued

```
5821 AAGAAATATTTTAATGCAGGTCATTCAGATGTAGCGGATAATGGAACTCTTTCTTAGGC
       TTCTTTATAAAATTACGTCCAGTAAGTCTACATCGCCTATTACCTTGAGAAAGAATCCG
         > K  K  Y  F  N  A  G  H  S  D  V  A  D  N  G  T  L  F  L  G

5881 ATTTTGAAGAATTGGAAAGAGGAGAGTGACAGAAAAATAATGCAGAGCCAAATTGTCTCC
       TAAAACTTCTTAACCTTTCTCCTCTCACTGTCTTTTTATTACGTCTCGGTTTAACAGAGG
         > I  L  K  N  W  K  E  E  S  D  R  K  I  M  Q  S  Q  I  V  S

5941 TTTTACTTCAAACTTTTTAAAAACTTTAAAGATGACCAGAGCATCCAAAAGAGTGTGGAG
       AAAATGAAGTTTGAAAAATTTTTGAAATTTCTACTGGTCTCGTAGGTTTTCTCACACCTC
         > F  Y  F  K  L  F  K  N  F  K  D  D  Q  S  I  Q  K  S  V  E

6001 ACCATCAAGGAAGACATGAATGTCAAGTTTTTCAATAGCAACAAAAAGAAACGAGATGAC
       TGGTAGTTCCTTCTGTACTTACAGTTCAAAAAGTTATCGTTGTTTTTCTTTGCTCTACTG
         > T  I  K  E  D  M  N  V  K  F  F  N  S  N  K  K  K  R  D  D

6061 TTCGAAAAGCTGACTAATTATTCGGTAACTGACTTGAATGTCCAACGCAAAGCAATACAT
       AAGCTTTTCGACTGATTAATAAGCCATTGACTGAACTTACAGGTTGCGTTTCGTTATGTA
         > F  E  K  L  T  N  Y  S  V  T  D  L  N  V  Q  R  K  A  I  H

6121 GAACTCATCCAAGTGATGGCTGAACTGTCGCCAGCAGCTAAAACAGGGAAGCGAAAAAGG
       CTTGAGTAGGTTCACTACCGACTTGACAGCGGTCGTCGATTTTGTCCCTTCGCTTTTTCC
         > E  L  I  Q  V  M  A  E  L  S  P  A  A  K  T  G  K  R  K  R

6181 AGTCAGATGCTGTTTCGAGGTCGAAGAGCATCCCAGTAAtggttgtcctgcctgcaatat
       TCAGTCTACGACAAAGCTCCAGCTTCTCGTAGGGTCATTaccaacaggacggacgttata
         > S  Q  M  L  F  R  G  R  R  A  S  Q  •

6241 ttgaatttaaatctaaatctatttattaatatttaacattatttatatgggaatatat
```

Figure 3 continued

```
                                                    aacttaaaatttagatttagataaataattattataaattgtaataatacccccttatata

6301 ttttagactcatcaatcaaataagtatttataatagcaacttTTTTGTaatggatccc
     aaaatctgagtagtttagtttattcataaatattatcgtttgaAAAAACAttacctaggg
                    Engineered transcriptional stop motif (in upper case)

6359 agctctctcgacgcaggactcggcttgctgaagcgcgcacagcaagaggcgagggcggc
     tcgagagagctgcgtcctgagccgaacgacttcgcgcgtgtcgttctcccgctcccgccg

6419 gactggtgagtacgccaatttttgactagcggaggctagaaggagagagAGGGTGCG
     ctgaccactcatgcggttaaaaactgatcgcctcccgatctttcctctctctcTACCCACGC
                                                    > M G A
                                    HIV gag protein coding sequence →

6479 AGAGCGTCGGTATTAAGCGGGGGAGAATTAGATAAATGGGAAAAAATTCGGTTAAGGCCA
     TCTCGCAGCCATAATTCGCCCCCTCTTAATCTATTTACCCTTTTTTAAGCCAATTCCGGT
       > R A S V L S G G E L D K W E K I R L R P

6539 GGGGGAAAGAAAAAATATAAGTTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGA
     CCCCCTTTCTTTTTTATATTCAATTTTGTATATCATACCCGTTCGTCCCTCGATCTTGCT
       > G G K K K Y K L K H I V W A S R E L E R

6599 TTCGCAGTCAATCCTGGCCTGTTAGAAACATCAGAAGGCTGCAGACAAATATTGGGACAG
     AAGCGTCAGTTAGGACCGGACAATCTTTGTAGTCTTCCGACGTCTGTTTATAACCCTGTC
       > F A V N P G L L E T S E G C R Q I L G Q

6659 CTACAGCCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCA
     GATGTCGGTAGGGAAGTCTGTCCTAGTCTTCTTGAATCTAGTAATATATTATGTCATCGT
       > L Q P S L Q T G S E E L R S L Y N T V A
```

Figure 3 continued

```
6719 ACCCTCTATTGTGTACATCAAAGGATAGATGTAAAAGACACCAAGGAAGCTTTAGAGAAG
       TGGGAGATAACACATGTAGTTTCCTATCTACATTTTCTGTGGTTCCTTCGAAATCTCTTC
         > T  L  Y  Y  C  V  H  Q  R  I  D  D  V  K  K  D  T  K  E  E  A  L  E  E  K

6779 ATAGAGGAAGAGAGCAAAACAAAAGTAAGAAAAAGGCACAGCAAGCAGCAGCTGCAGCTGGC
       TATCTCCTTCTCGTTTTGTTTTCATTCTTTTTCCGTGTCGTTCGTCGACGTCGACCG
         > I  E  E  E  Q  N  K  S  K  K  K  A  Q  Q  Q  A  A  A  A  A  G

6839 ACAGGAAACAGCAGCCAGGTCAGCCAAAATTACCCTATAGTGCAGAACCTACAGGGGCAA
       TGTCCTTTGTCGTCGGTCCAGTCGGTTTTAATGGGATATCACGTCTTGGATGTCCCCGTT
         > T  G  N  S  S  Q  V  S  Q  N  Y  P  I  V  Q  N  L  Q  G  Q

6899 ATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAAGTAGTAGAAGAA
       TACCATGTAGTCCGGTATAGTGGATCTTGAAATTTACGTACCCATTTTCATCATCTTCTT
         > M  V  H  Q  A  I  S  P  R  T  L  N  A  W  V  K  V  V  E  E

6959 AAGGCTTTCAGCCCAGAAGTAATACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCA
       TTCCGAAAGTCGGGTCTTCATTATGGGTACAAAAGTCGTAATAGTCTTCCTCGGTGGGGT
         > K  A  F  S  P  E  V  I  P  M  F  S  A  L  S  E  G  A  T  P

7019 CAAGATTTAAACACCATGCTAAACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTA
       GTTCTAAATTTGTGGTACGATTTGTGTCACCCCCCTGTAGTTCGTCGGTACGTTTACAAT
         > Q  D  L  N  T  M  L  N  T  V  G  G  H  Q  A  A  M  Q  M  L

7079 AAAGAGACTATCAATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGG
       TTTCTCTGATAGTTACTCCTTCGACGTCTTACCCTATCTCACGTAGGTCACGTACGTCCC
         > K  E  T  I  N  E  E  A  A  E  W  D  R  V  H  P  V  H  A  G
```

Figure 3 continued

```
7139 CCTATTGCACCAGGCCAAATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACTACTAGT
       GGATAACGTGGTCCGGTTTACTCTCTTGGTTCCCCTTCACTGTATCGTCCTTGATGATCA
         > P  I  A  P  G  Q  M  R  E  P  R  G  S  D  I  A  G  T  T  S

7199 ACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCTATCCCAGTAGGAGAAATC
       TGGGAAGTCCTTGTTTATCCTACCTACTGTTTATTAGGTGGATAGGGTCATCCTCTTTAG
         > T  L  Q  E  Q  I  G  W  M  T  N  N  P  P  I  P  V  G  E  I

7259 TATAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGC
       ATATTTTCTACCTATTAGGACCCTAATTTATTTTATCATTCTTACATATCGGGATGGTCG
         > Y  K  R  W  I  I  L  G  L  N  K  I  V  R  M  Y  S  P  T  S

7319 ATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGATTATGTAGACCGGTTCTAT
       TAAGACCTGTATTCTGTTCCTGGTTTCCTTGGGAAATCTCTAATACATCTGGCCAAGATA
         > I  L  D  I  R  Q  G  P  K  E  P  F  R  D  Y  V  D  R  F  Y

7379 AAAACTCTAAGAGCCGAACAAGCTTCACAGGATGTAAAAAATTGGATGACAGAAACCTTG
       TTTTGAGATTCTCGGCTTGTTCGAAGTGTCCTACATTTTTTAACCTACTGTCTTTGGAAC
         > K  T  L  R  A  E  Q  A  S  Q  D  D  V  K  K  N  W  M  T  E  T  L

7439 TTGGTCCAAAATGCAAACCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCAGCT
       AACCAGGTTTTACGTTTGGGTCTAACATTCTGATAAAATTTTCGTAACCCTGGTCGTCGA
         > L  V  Q  N  A  N  P  D  D  C  K  T  I  L  L  K  A  L  G  P  A  A

7499 ACACTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGGGGACCCGGCCATAAAGCAAGA
       TGTGATCTTCTTTACTACTGTCGTACAGTCCCTCACCCCCCTGGGCCGGTATTTCGTTCT
         > T  L  E  E  M  M  T  A  C  Q  G  V  G  G  G  P  G  H  K  A  R

7559 GTTTTGGCTGAAGCCATGAGCCAAGTAACAAATCCAGCTAACATAATGATGCAGAGAGGC
```

Figure 3 continued

```
                                                    CAAAACCGACTTCGGTACTCGGTTCATTGTTTAGGTCGATTGTATTACTACGTCTCTCCG
                                                      > V  L  A  E  A  M  S  Q  V  T  N  P  A  N  I  M  M  Q  R  G

7619 AATTTTAGGAACCAAAGAAAGACTGTTAAGTGTTTCAATTGTGGCAAAGAAGGGCACATA
       TTAAAATCCTTGGTTTCTTTCTGACAATTCACAAAGTTAACACCGTTTCTTCCCGTGTAT
         > N  F  R  N  Q  R  K  T  V  K  K  C  F  N  C  G  K  E  G  H  I

7679 GCCAAAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAGATGTGGAAGGGAAGGACAC
       CGGTTTTTAACGTCCCGGGGATCCTTTTTCCCGACAACCTCTACACCTTCCCTTCCTGTG
         > A  K  N  C  R  A  P  R  K  K  G  C  W  R  C  G  R  E  G  H

7739 CAAATGAAAGATTGCACTGAGAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCCTAC
       GTTTACTTTCTAACGTGACTCTCTGTCCGATTAAAAAATCCCTTCTAGACCGGAAGGATG
         > Q  M  K  D  C  T  E  R  Q  A  N  F  L  G  K  I  W  P  S  Y
                                                    > F  F  R  E  D  L  A  F  L
                                                    HIV pol protein coding sequence →

7799 AAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGAG
       TTCCCTTCCGGTCCCTTAAAAGAAGTCTCGTCTGGTCTCGGTTGTCGGGGTGGTCTTCTC
         > K  G  R  P  G  N  F  L  Q  S  R  P  E  P  T  A  P  P  E  E
         >Q  G  K  A  R  E  F  S  S  E  Q  T  R  A  N  S  P  T  R  R

7859 AGCTTCAGGTTTGGGGAGAGAAAACAACTCCCTCTCAGAAGCAGGAGCCGATAGACAAG
       TCGAAGTCCAAACCCCTCCTCTTTGTTGAGGGAGAGTCTTCGTCCTCGGCTATCTGTTC
         > S  F  R  F  G  E  E  K  T  T  P  S  Q  K  Q  E  P  I  D  K
         >E  L  Q  V  W  G  G  E  N  N  S  L  S  E  A  G  A  D  R  Q

7919 GAACTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAATAA
       CTTGACATAGGAAATTGAAGGGAGTCTAGTGAGAAACCGTTGCTGGGGAGCAGTGTTATT
```

Figure 3 continued

```
      >E L Y P L T S L R S L F G N D P S S Q *
      >G T V S F N F P Q I T L W Q R P L V T I

7979  GGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAG
        CCTATCCCCCCGTTGATTTCCTTCGAGATAATCTATGTCCTCGTCTACTATGTCATAATC
          >R I G G Q L K E A L L D T G A D D T V L

8039  AAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTA
        TTCTTTACTTAAACGGTCCTTTTACCTTTGGTTTTTACTATCCCCCTTAACCTCCAAAAT
          >E E M N L P G K W K P K M I G G I G G F

8099  TCAAAGTAAGACAGTACGATCAGATACCTGTAGAAATCTGTGGACATAAAGCTATAGGTA
        AGTTTCATTCTGTCATGCTAGTCTATGGACATCTTTAGACACCTGTATTTCGATATCCAT
          >I K V R Q Y D Q I P V E I C G H K A I G

8159  CAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTG
        GTCATAATCATCCTGGATGTGGACAGTTGTATTAACCTTCTTTAGACAACTGAGTCTAAC
          >T V L V G P T P V N I I G R N L L T Q I

8219  GTTGTACTTTAAATTTCCCCATTAGTCCTATTGAAACTGTACCAGTAAAATTAAAGCCAG
        CAACATGAAATTTAAAGGGGTAATCAGGATAACTTTGACATGGTCATTTTAATTTCGGTC
          >G C T L N F P I S P I E T V P V K L K P

8279  GAATGGATGGCCCAAAAGTTAAGCAATGGCCATTGACAGAAGAAAAATAAAAGCATTAG
        CTTACCTACCGGGTTTTCAATTCGTTACCGGTAACTGTCTTCTTTTTATTTTCGTAATC
          >G M D G P K V K Q W P L T E E K I K A L

8339  TAGAGATATGTACAGAAATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATC
        ATCTCTATACATGTCTTTACCTTTTCCTTCCCTTTTAAGTTTTTTAACCCGGACTTTTAG
```

Figure 3 continued

```
      >V  E  I  C  T  E  M  E  K  E  G  K  K  I  S  K  I  G  P  E  N

8399  CATACAATACTCCAGTATTTGCTATAAAGAAAAAAGACAGTACTAAATGGAGAAAACTAG
        GTATGTTATGAGGTCATAAACGATATTTCTTTTTTCTGTCATGATTTACCTCTTTTGATC
        >P  Y  N  T  P  V  F  A  I  K  K  K  K  D  S  T  K  W  R  K  L

8459  TAGATTTCAGAGAACTTAATAAAAGAACTCAAGACTTCTGGGAAGTTCAGTTAGGAATAC
        ATCTAAAGTCTCTTGAATTATTTTCTTGAGTTCTGAAGACCCTTCAAGTCAATCCTTATG
        >V  D  F  R  E  L  N  K  R  T  Q  D  F  W  E  V  Q  L  G  I

8519  CACACCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTATTGGATGTGGGTGATGCAT
        GTGTGGGGCGTCCCAATTTTTCTTTTTTAGTCATTGTCATAACCTACACCCACTACGTA
        >P  H  P  A  G  L  K  K  K  K  S  V  T  V  L  D  V  G  D  A

8579  ACTTTTCAGTTCCCTTAGATAAAGACTTTAGAAAGTATACTGCATTTACCATACCTAGTA
        TGAAAAGTCAAGGGAATCTATTTCTGAAATCTTTCATATGACGTAAATGGTATGGATCAT
        >Y  F  S  V  P  L  D  K  D  F  R  K  K  Y  T  A  F  T  I  P  S

8639  TAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTGCCACAGGGATGGAAAG
        ATTTGTTACTCTGTGGTCCCTAATCTATAGTCATGTTACACGACGGTGTCCCTACCTTTC
        >I  N  N  E  T  P  G  I  R  Y  Q  Y  N  V  L  P  Q  G  W  K

8699  GATCACCAGCAATATTCCAAAGTAGCATGACACAAAAATCTTAGAGCCTTTTAGAAAACAGA
        CTAGTGGTCGTTATAAGGTTTCATCGTACTGTTTTTAGAATCTCGGAAAATCTTTTGTCT
        >G  S  P  A  I  F  Q  S  S  M  T  K  I  L  L  E  P  F  R  K  Q

8759  ATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAA
        TAGGTCTGTATCAATAGATAGTTATGTACCTACTAAACATACATCCTAGACTGAATCTTT
        >N  P  D  I  V  I  Y  Q  Y  M  D  D  L  Y  V  G  S  D  L  E
```

Figure 3 continued

```
8819 TAGGGCAGCATAGAACAAAAATAGAGGAACTGAGACAGCATCTGTTGAGGTGGGGATTTA
       ATCCCGTCGTATCTTGTTTTTATCTCCTTGACTCTGTCGTAGACAACTCCACCCCTAAAT
         >I  G  Q  H  R  T  K  I  E  E  L  R  Q  H  L  L  R  W  G  F

8879 CCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCC
       GGTGTGGTCTGTTTTTTGTAGTCTTTCTTGGAGGTAAGGAAACCTACCCAATACTTGAGG
         >T  T  P  D  K  K  H  Q  K  E  P  P  F  L  W  M  G  Y  E  L

8939 ATCCTGATAAATGGACAGTACAGCCTATAATGCTGCCAGAAAAAGACAGCTGGACTGTCA
       TAGGACTATTTACCTGTCATGTCGGATATTACGACGGTCTTTTTCTGTCGACCTGACAGT
         >H  P  D  K  W  T  V  Q  P  I  M  L  P  E  K  D  S  W  T  V

8999 ATGACATACAGAAGTTAGTGGGAAAATTGAATTGGGCAAGTCAGATTTATGCAGGGATTA
       TACTGTATGTCTTCAATCACCCTTTTAACTTAACCCGTTCAGTCTAAATACGTCCCTAAT
         >N  D  I  Q  K  L  V  G  K  L  N  W  A  S  Q  I  Y  A  G  I

9059 AAGTAAAGCAGTTATGTAAACTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCAC
       TTCATTTCGTCAATACATTTGAGGAATCTCCTTGGTTTCGTGATTGTCTTCATTATGGTG
         >K  V  K  Q  L  C  K  L  L  R  G  T  K  A  L  T  E  V  I  P

9119 TAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGGGAGATTCTAAAAGAACCAGTAC
       ATTGTCTTCTTCGTCTCGATCTTGACCGTCTTTTGTCCCTCTAAGATTTTCTTGGTCATG
         >L  T  E  E  A  E  L  E  L  A  E  N  R  E  I  L  K  E  P  V

9179 ATGAAGTATATTATGACCCATCAAAAGACTTAGTAGCAGAAATACAGAAGCAGGGCAAG
       TACTTCATATAATACTGGGTAGTTTTCTGAATCATCGTCTTTATGTCTTCGTCCCCGTTC
         >H  E  V  Y  Y  D  P  S  K  D  L  V  A  E  I  Q  K  Q  G  Q
```

Figure 3 continued

```
9239 GCCAATGGACATATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAGTATG
     CGGTTACCTGTATAGTTTAAATAGTTCTCGGTAAATTTTTAGACTTTTGTCCTTTCATAC
       >G  Q  W  T  Y  Y  Q  I  I  Y  Y  Q  E  P  F  K  K  N  L  L  K  T  G  K  K  Y

9299 CAAGGATGAGGGGTGCCCACACTAATGATGTAAAACAGTTAACAGAGGCAGTGCAAAAAG
       GTTCCTACTCCCCCACGGGTGTGATTACTACATTTTGTCAATTGTCTCCGTCACGTTTTTC
       >A  R  M  R  G  A  H  T  N  D  V  K  Q  L  T  E  A  V  Q  K

9359 TATCCACAGAAAGCATAGTAATATGGGGAAAGATTCCTAAATTTAAACTACCCATACAAA
       ATAGGTGTCTTTCGTATCATTATACCCCTTTCTAAGGATTTAAATTTGATGGGTATGTTT
       >V  S  T  E  S  I  V  I  W  G  K  I  P  K  F  K  L  P  I  Q

9419 AGGAAACATGGGAAGCATGGTGGATGGAGTATTGGCAAGCTACCTGGATTCCTGAGTGGG
       TCCTTTGTACCCTTCGTACCACCTACCTCATAACCGTTCGATGGACCTAAGGACTCACCC
       >K  E  T  W  E  A  W  W  M  E  Y  W  Q  A  T  W  I  P  E  W

9479 AGTTTGTCAATACCCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAGAACCCATAG
       TCAAACAGTTATGGGGAGGGAATCACTTTAATACCATGGTCAATCTCTTTCTTGGGTATC
       >E  F  V  N  T  P  P  L  V  K  L  W  Y  Q  L  E  K  E  P  I

9539 TAGGAGCAGAAACTTTCTATGTAGATGGGGCAGCTAATAGGGAGACTAAATTAGGAAAAG
       ATCCTCGTCTTTGAAAGATACATCTACCCCGTCGATTATCCCTCTGATTTAATCCTTTTC
       >V  G  A  E  T  F  Y  V  D  G  A  A  N  R  E  T  K  L  G  K

9599 CAGGATATGTTACTGACAGAGGAAGACAAAAAGTTGTCTCCATAGCTGACACAACAAATC
       GTCCTATACAATGACTGTCTCCTTCTGTTTTTCAACAGAGGTATCGACTGTGTTGTTTAG
       >A  G  Y  Y  V  T  D  R  G  R  Q  K  V  V  S  I  A  D  T  T  N

9659 AGAAGACTGAATTACAAGCAATTCATCTAGCTTTGCAGGATTCGGGATTAGAAGTAAACA
```

Figure 3 continued

```
      TCTTCTGACTTAATGTTCGTTAAGTAGATCGAAACGTCCTAAGCCCTAATCTTCATTTGT
        >Q  K  T  E  L  Q  A  I  H  L  A  L  Q  D  S  G  L  E  V  N

9719  TAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGATAAGAGTGAAT
        ATCATTGTCTGAGTGTTATACGTAATCCTTAGTAAGTTCGTGTTGGTCTATTCTCACTTA
          >I  V  T  D  S  Q  Y  A  L  G  I  I  Q  A  Q  P  D  K  S  E

9779  CAGAGTTAGTCAGTCAAATAATAGAGCAGTTAATAAAAAAGGAAAGGTCTACCTGGCAT
        GTCTCAATCAGTCAGTTTATTATCTCGTCAATTATTTTTCCTTTTCCAGATGGACCGTA
          >S  E  L  V  S  Q  I  I  E  Q  L  I  K  K  E  K  V  Y  L  A

9839  GGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGCTG
        CCCATGGTCGTGTGTTTCCTTAACCTCCTTTACTTGTTCATCTATTTAATCAGTCACGAC
          >W  V  P  A  H  K  G  I  G  G  N  E  Q  V  D  K  L  V  S  A

9899  GAATCAGGAAAGTACTATTTTTGAATGGAATAGATAAGGCCCAAGAAGAACATGAGAAAT
        CTTAGTCCTTTCATGATAAAAACTTACCTTATCTATTCCGGGTTCTTCTTGTACTCTTTA
          >G  I  R  K  V  L  F  L  N  G  I  D  K  A  Q  E  E  H  E  K

9959  ATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTGTAGTAGCAAAAG
        TAGTGTCATTAACCTCTCGTTACCGATCACTAAAATTGGACGGTGGACATCATCGTTTTC
          >Y  H  S  N  W  R  A  M  A  S  D  F  N  L  P  P  V  V  A  K

10019 AAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAG
        TTTATCATCGGTCGACACTATTTACAGTCGATTTTCCTCTTCGGTACGTACCTGTTCATC
          >E  I  V  A  S  C  D  K  C  Q  L  K  G  E  A  M  H  G  Q  V

10079 ACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATCTAGAAGGAAAAATTATCCTGG
        TGACATCAGGTCCTTATACCGTTGATCTAACATGTGTAGATCTTCCTTTTTAATAGGACC
```

Figure 3 continued

```
      >D  C  S  P  G  I  W  Q  L  D  C  T  H  L  E  G  K  I  I  L
10139 TAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTTATTCCAGCAGAGACAGGGC
        ATCGTCAAGTACATCGGTCACCTATATATCTTCGTCTTCAATAAGGTCGTCTCTGTCCCG
          >V  A  V  H  V  A  S  G  Y  I  E  A  E  V  I  P  A  E  T  G
10199 AGGAAACAGCATATTTTCTCTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAATACATA
        TCCTTTGTCGTATAAAAGAGAATTTTAATCGTCCTTCTACCGGTCATTTTTGTTATGTAT
          >Q  E  T  A  Y  F  L  L  K  L  A  G  R  W  P  V  K  T  I  H
10259 CAGACAATGGCAGCAATTTCACCAGTACTACGGTTAAGGCCGCCTGTTGGTGGGCAGGGA
        GTCTGTTACCGTCGTTAAAGTGGTCATGATGCCAATTCCGGCGGACAACCACCCGTCCCT
          >T  D  N  G  S  N  F  T  S  T  T  V  K  A  A  C  W  W  A  G
10319 TCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGA
        AGTTCGTCCTTAAACCGTAAGGGATGTTAGGGGTTTCAGTTCCTCATCATCTTAGATACT
          >I  K  Q  E  F  G  I  P  Y  N  P  Q  S  Q  G  V  V  E  S  M
10379 ATAATGAATTAAAGAAAATTATAGGACAGGTAAGAGATCAGGCTGAACACCTTAAGACAG
        TATTACTTAATTTCTTTTAATATCCTGTCCATTCTCTAGTCCGACTTGTGGAATTCTGTC
          >N  N  E  L  K  K  I  I  G  Q  V  R  D  Q  A  E  H  L  K  T
10439 CAGTACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGATACA
        GTCATGTTTACCGTCATAAGTAGGTGTTAAAATTTTCTTTTCCCCCCTAACCCCCTATGT
          >A  V  Q  M  A  V  F  I  H  N  F  K  R  K  G  G  I  G  G  Y
10499 GTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAACTACAAA
        CACGTCCCCTTTCTTATCATCTGTATTATCGTTGTCTGTATGTTTGATTTCTTGATGTTT
          >S  A  G  E  R  I  V  D  I  I  A  T  D  I  Q  T  K  E  L  Q
```

Figure 3 continued

```
10559 AGCAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAACAAAGATCCCCTTT
        TCGTTTAATGTTTTTAAAAGCCCAAATAATGTCCCTGTTGTTTCTAGGGGAAA
          >K  Q  I  T  K  I  Q  N  F  R  R  V  V  Y  Y  Y  R  D  N  K  D  P  L

10619 GGAAAGGACCAGCAAAGCTTCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATA
        CCTTTCCTGGTCGTTTCGAAGAGACCTTTCCCCGTCATCATTATGTTCTATTAT
          >W  K  G  P  A  K  L  L  W  K  G  G  E  G  A  V  V  I  Q  D  N

10679 GTGACATAAAAGTAGTGCCAAGAAGAAAAGCAAAAATCATTAGGGATTATGGAAAACAGA
        CACTGTATTTTCATCACGGTTCTTCTTTTCGTTTTAGTAATCCCTAATACCTTTTGTCT
          >S  D  I  K  V  V  P  R  R  K  A  K  I  I  R  R  D  Y  G  K  Q

10739 TGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAGaacatggaaaagttta
        ACCGTCCACTACTAACACACCGTTCATCTGTCCTACTCCTAATCttgtacctttttcaaat
          >M  A  G  D  D  C  V  A  S  R  Q  D  E  D  •

10799 gtaaaacaccatagggtcgactgcagaagcttccatgggagctctttagtgtaataaat
        catttgtggtatcccagctgacgtcttcgaaggtacccctcgagaaatcacattattta

10859 ttaataaaatattgacaaaatagttaaatgaatatatgaaagtacattatacacggaATG
        aattattttataactgttttatcaatttacttttcatgtaatatgtgccttAC

10919 GAGTTCGATATTAGTTCTTGCAGAATGATATATTCTGTTCTCGAACAATATCACTTTGTT
        CTCAAGCTATAATCAAGAACGTCTTACTATATAAGACAAGAGCTTGTTATAGTGAAACAA

10979 ACTGATAATCGTTATAACAACCATAATCAAAAATTTAGAATTATATTATACTGTTTAAAA
        TGACTATTAGCAATATTGTTGGTATTAGTTTTTAAATCTTAATATATGACAAATTTT
Fowlpox virus 3' flanking region of insertion site (in upper case) →
```

Figure 3 continued

```
11039 GATTCTACGATAAAGAAATATCCGTACAGGTTTGTTTCTGAAATTCACTTTGTAAGATAC
      CTAAGATGCTATTTCTTTATAGGCATGTCCAAACAAAGACTTTAAGTGAAACATTCTATG

11099 ATAATTAACAAATTCAGGGGGAAAAATCTTTACAAAATTAGTATAGAAGCTATAGATATA
      TATTAATTGTTTAAGTCCCCCTTTTTAGAAATGTTTTAATCATATCTTCGATATCTATAT

11159 TCAAAAGGTAGACAACAAATAATCAGAACCTAATTTTTTTATCAAAAATTAAAATATAA
      AGTTTTCCATCTGTTGTTTATTAGTCTTGGATTAAAAAAATAGTTTTTTTAATTTTATATT

11219 ATAAAATGAAAAATAACTTGTATGAAGAAAAAATGAACATGAGTAAGAAACAAGTAAAAA
      TATTTTACTTTTTTATTGAACATACTTCTTTTTTTACTTGTACTCATTCTTTGTTCATTTTT

11279 CTCAAAGTAAATGTAATAATAACGCATCTAGATTTACATGCCTGGATGCGGTGCA
      GAGTTTCATTTACATTATTATTGCGTAGATCTAAATGTACGGACCTACGCCACGT
```

Figure 3 continued